(a) selecting a frame cluster in said input video sequence which corresponds to a most static one of said video segments;

(b) computing a content value in said selected frame cluster;

(c) using said computed content value to cluster remaining frames in said input video sequence.

2. The method of claim 1, wherein in said (a) said frame cluster is selected using a refined feature space representation of said input video sequence.

3. The method of claim 1, wherein in said (a) each of said plurality of frames is transformed into a histogram vector indicative of a spatial distribution of colors in said each of said plurality of frames.

4. The method of claim 3, wherein in said (a) each of said plurality of frames is divided into a plurality of blocks, each of said plurality of blocks being represented by a histogram in a color space indicative of a distribution of colors within each of said plurality of blocks.

5. The method of claim 3, wherein each of said plurality of frames is divided into a plurality of blocks and each said histogram vector comprises a plurality of histograms in a color space, each of said plurality of histograms corresponding to one of said plurality of blocks.

6. The method of claim 2, wherein said refined feature space representation is obtained using a singular value decomposition of said input video sequence.

7. The method of claim 6, wherein said singular value decomposition is performed using frames selected with a fixed interval from said input video sequence.

8. The method of claim 7, wherein said selected frames are arranged into a feature frame matrix, and wherein said singular value decomposition is performed on said feature frame matrix.

9. The method of claim 6, wherein said singular value decomposition produces a matrix, each column of said matrix representing a frame in a refined feature space corresponding to a frame in said input video sequence.

10. The method of claim 1, further comprising (d) using said clustered frames to output a motion video representative of a summary of said input video sequence.

11. The method of claim 1, further comprising (d) outputting a plurality of keyframes, each of said plurality of keyframes representative of said clustered frames.

12. The method of claim 2, wherein said selecting comprises locating a cluster closest to an origin of said refined feature space.

13. The method of claim 2, wherein said (c) comprises:

(c)(1) sorting a plurality of vectors in said refined feature space in ascending order according to a distance of each of said vectors to an origin of said refined feature space representation;

(c)(2) selecting a vector among said sorted vectors which is closest to an origin of said refined feature space representation and including said selected vector into a first cluster;

(c)(3) clustering said plurality of sorted vectors in said refined feature into a plurality of clusters according to a distance between each of said plurality of sorted vectors and vectors in each of said plurality of clusters and an amount of information in each of said plurality of clusters.

14. The method of claim 13, wherein in said (c)(3) said plurality of sorted vectors are clustered into said plurality of clusters such that said amount of information in each of said plurality of clusters does not exceed an amount of information in said first cluster.

15. The method of claim 13, wherein said first cluster is composed of frames based on a distance variation between said frames and an average distance between frames in said first cluster.

16. The method of claim 13, wherein each of said plurality of clusters is composed of frames based on a distance variation between said frames and an average distance between frames in said each of said plurality of clusters.

17. A method for summarizing a content of an input video sequence, said method comprising:

(a) selecting frames from said input video sequence, said selected frames being taken at a fixed interval;

(b) creating a feature frame matrix using said selected frames;

(c) performing a singular value decomposition on said feature frame matrix to obtain a matrix representing said video sequence in a refined feature space;

(d) selecting a cluster in said refined feature space corresponding to a most static video segment;

(e) computing a content value corresponding to said selected cluster;

(f) using said computed content value to cluster frames in said input video sequence.

18. A method for segmenting an input video sequence, said input video sequence comprising a plurality of frames, said plurality of frames being grouped into a plurality of video shots, said method comprising:

(a) computing a similarity between each of said plurality of frames and a frame preceding said each of said plurality of frames in time;

(b) segmenting said input video sequence into said plurality of video shots according to said computed similarity.

19. The method of claim 18, wherein said similarity is calculated using a refined feature space representation of said input video sequence.

20. The method of claim 19, wherein said refined feature space representation is created using a singular value decomposition of said input video sequence.

21. The method of claim 20, wherein said singular value decomposition is performed using frames selected with a fixed interval from said input video sequence.

22. The method of claim 21, wherein said selected frames are arranged into a feature frame matrix, and wherein said singular value decomposition is performed on said feature frame matrix.

23. The method of claim 22, wherein said performed singular value decomposition produces a matrix, each column of said produced matrix comprising a frame in said refined feature space representing a frame in said input video sequence.

24. The method of claim 18, further comprising (c) extracting features from each of said plurality of video shots.

25. A method for determining a similarity between a first and a second frame in an input video sequence, said method comprising:

(a) calculating a refined feature space representation of said input video sequence;

(b) using said calculated representation to compute said similarity between said first and said second frames.

26. The method of claim 25, wherein in said (a) said refined feature space representation is calculated using a singular value decomposition.

27. The method of claim 18, wherein in said (b) said computed similarity is compared to at least a first threshold similarity and a second threshold similarity, and said input video sequence is segmented according to a result of said comparison.

28. The method of claim 18, wherein if in said (b) said computed similarity is below a first threshold similarity, said each of said plurality of frames is put into a one of said plurality of video shots containing said precedent in time frame.

29. The method of claim 18, wherein if in said (b) said computed similarity is above a second threshold similarity, said each of said plurality of frames is designated as a shot boundary.

30. The method of claim 18, wherein if in said (b) said computed similarity is between a first threshold similarity and a second threshold similarity, said each of said plurality of frames is put into a one of said plurality of video shots according to a further analysis performed using additional frames from said plurality of frames.

sub. B1

31. A computer-readable medium containing a program for summarizing a content of an input video sequence, said input video sequence comprising a plurality of frames, said plurality of frames being grouped into a plurality of video segments, said program comprising:

(a) selecting a frame cluster in said input video sequence which corresponds to a most static video segment;

(b) computing content value in said selected frame cluster;

(c) using said computed content value to cluster remaining frames in said input video sequence.

32. The computer-readable medium of claim 31, wherein in said (a) said frame cluster is selected using a refined feature space representation of said input video sequence.

33. The computer-readable medium of claim 31, wherein in said (a) each of said plurality of frames is transformed into a histogram vector indicative of a spatial distribution of colors in said each of said plurality of frames.

34. The computer-readable medium of claim 33, wherein in said (a) each of said plurality of frames is divided into a plurality of blocks, each of said plurality of blocks being represented by a histogram in a color space indicative of a distribution of colors within each of said plurality of blocks.

35. The computer-readable medium of claim 33, wherein each of said plurality of frames is divided into a plurality of blocks and each said histogram vector comprises a plurality of histograms in a color space, each of said plurality of histograms corresponding to one of said plurality of blocks.

36. The computer-readable medium of claim 32, wherein said refined feature space representation is obtained using a singular value decomposition of said input video sequence.

37. The computer-readable medium of claim 36, wherein said singular value decomposition is performed using frames selected with a fixed interval from said input video sequence.

38. The computer-readable medium of claim 37, wherein said selected frames are arranged into a feature frame matrix, and wherein said singular value decomposition is performed on said feature frame matrix.

39. The computer-readable medium of claim 33, wherein said singular value decomposition produces a matrix, each column of said matrix representing a frame in a refined feature space corresponding to a frame in said input video sequence.

40. The computer-readable medium of claim 31, further comprising (d) using said clustered frames to output a video representative of a summary of said input video sequence.

41. The computer-readable medium of claim 31, further comprising (d) outputting a plurality of keyframes, each of said plurality of keyframes representative of said clustered frames.

42. The computer-readable medium of claim 32, wherein said selecting comprises locating a cluster closest to an origin of said refined feature space.

43. The computer-readable medium of claim 32, wherein said (c) comprises:

(1) sorting a plurality of vectors in said refined feature space in ascending order according to a distance of each of said vectors to an origin of said refined feature space;

(2) selecting a vector among said sorted vectors which is closest to an origin of said refined feature space and including said selected vector into a first cluster;

(3) clustering said plurality of sorted vectors in said refined feature into a plurality of clusters according to a distance between each of said plurality of sorted vectors and each of said plurality of clusters and an amount of information in each of said plurality of clusters.

44. The computer-readable medium of claim 38, wherein in said (3) said plurality of sorted vectors are clustered into said plurality of clusters such that said amount of information in each of said plurality of clusters does not exceed an amount of information in said first cluster.

45. The computer-readable medium of claim 38, wherein said first cluster is composed of frames based on a distance variation between said frames and said first cluster.

46. The computer-readable medium of claim 38, wherein each of said plurality of clusters is composed of frames based on a distance variation between said frames and said each of said plurality of clusters.

47. A computer-readable medium containing a program for summarizing a content of an input video sequence, said program comprising:

(a) selecting frames with a fixed interval from said input video sequence;

(b) creating a feature frame matrix using said selected frames;

(c) performing a singular value decomposition on said feature frame matrix to obtain matrix representing said video sequence in refined feature space;

(d) selecting a cluster in said refined feature space corresponding to a most static video segment;

(e) computing a content value corresponding to said selected cluster;

(f) using said computed content value to cluster frames in said input video sequence.

48. A computer-readable medium containing a program for segmenting an input video sequence, said input video sequence comprising a plurality of frames, said plurality of frames being grouped into a plurality of video shots, said program comprising:

(a) computing a similarity between each of said plurality of frames and a subsequent in time frame;

(b) segmenting said input video sequence into a plurality of shots according to said computed similarity.

49. The computer-readable medium of claim 18, wherein said similarity is calculated using a refined feature space representation of said input video sequence.

50. The computer-readable medium of claim 19, wherein said refined feature space representation is created using a singular value decomposition of said input video sequence.

51. The computer-readable medium of claim 20, wherein said singular value decomposition is performed using frames selected with a fixed interval from said input video sequence.

52. The computer-readable medium of claim 21, wherein said selected frames are arranged into a feature frame matrix, and wherein said singular value decomposition is performed on said feature frame matrix.

53. The computer-readable medium of claim 22, wherein said performed singular value decomposition produces a matrix, each column of said produced matrix comprising a frame in said refined feature space representing a frame in said input video sequence.

54. The computer-readable medium of claim 18, wherein said program further comprises (c) extracting features from each of said plurality of video shots.

55. A computer-readable medium containing a program for determining a similarity between a first and a second frames in an input video sequence, said program comprising:

(a) calculating a refined feature space representation of said input video sequence; and

(b) using said calculated representation to compute said similarity between said first and said second frames.

56. The computer-readable medium of claim 25, wherein in said (a) said refined feature space representation is calculated using a singular value decomposition.

57. The computer-readable medium of claim 18, wherein in said (b) said computed similarity is compared to at least two threshold similarities, and said input video sequence is segmented according to a result of said comparison.

58. The computer-readable medium of claim 48, wherein if in said (b) said computed similarity is below a first threshold similarity, said each of said plurality of frames is put into a one of said plurality of video shots containing said precedent in time frame.

59. The computer-readable medium of claim 48, wherein if in said (b) said computed similarity is above a second threshold similarity, said each of said plurality of frames is designated as a shot boundary.

60. The computer-readable medium of claim 48, wherein if in said (b) said computed similarity is between a first threshold similarity and a second threshold similarity, said each of said plurality of frames is put into a one of said plurality of video shots according to a further analysis performed using additional frames from said plurality of frames.

61. The method of claim 18, further comprising (c) extracting features from each of said plurality of video shots and using said extracted features to index said plurality of video shots.

62. The method of claim 61, wherein said extracted features are features of a video frame representative of said each of said plurality of video shots.

63. The computer-readable medium of claim 48, wherein said program further comprises (c) extracting features from each of said plurality of video shots and using said extracted features to index said plurality of video shots.

64. The method of claim 63, wherein said extracted features are features of a video frame representative of said each of said plurality of video shots.

65. A method of calculating a degree of visual changes in a video shot, said video shot comprising a plurality of frames, said method comprising:

(a) performing a singular value decomposition on said plurality of frames, wherein said singular value decomposition produces a matrix, each column of said matrix representing a frame in a refined feature space corresponding to a frame in said plurality of frames;

(b) using said matrix to calculate said degree of visual changes in said video shot.

66. The method of claim 65, wherein said (b) comprises calculating said degree of visual changes in said video shot as a sum $\sqrt{\sum_{j=1}^{rank(A)} V_{ij}^2}$, wherein v_{ij} are elements of said matrix.

67. A computer-readable medium containing a program for calculating a degree of visual changes in a video shot, said video shot comprising a plurality of frames, said program comprising:

(a) performing a singular value decomposition on said plurality of frames, wherein said singular value decomposition produces a matrix, each column of said matrix representing a frame in a refined feature space corresponding to a frame in said plurality of frames;

(b) using said matrix to calculate said degree of visual changes in said video shot.

68. The computer-readable medium of claim 67, wherein said (b) comprises calculating said degree of visual changes in said video shot as a sum $\sqrt{\sum_{j=1}^{rank(A)} V_{ij}^2}$, wherein v_{ij} are elements of said matrix.

69. A method of calculating an evenness of color distributions in a video shot, said video shot comprising a plurality of frames, said method comprising:

(a) performing a singular value decomposition on said plurality of frames, wherein said singular value decomposition produces a matrix, each column of said matrix representing a frame in a refined feature space corresponding to a frame in said plurality of frames;

(b) using said matrix to calculate said evenness of color distribution in said video shot.

70. The method of claim 69, wherein said (b) comprises calculating said evenness of color distribution in said video shot as a sum $\sqrt{\sum_{j=1}^{rank(A)} \sigma_j^2 \, v_{ij}^2}$, wherein said v_{ij} are elements of said matrix and said σ_j are singular values obtained in said singular value decomposition.

71. A computer-readable medium containing a program for calculating an evenness of color distributions in a video shot, said video shot comprising a plurality of frames, said method comprising:

(a) performing a singular value decomposition on said plurality of frames, wherein said singular value decomposition produces a matrix, each column of said matrix representing a frame in a refined feature space corresponding to a frame in said plurality of frames;

(b) using said matrix to calculate said evenness of color distribution in said video shot.

72. The computer readable medium of claim 71, wherein said (b) comprises calculating said evenness of color distribution in said video shot as a sum $\sqrt{\sum_{j=1}^{rank(A)} \sigma_j^2 \, v_{ij}^2}$, wherein said v_{ij} are elements of said matrix and said σ_j are singular values obtained in said singular value decomposition.